ARTICLE

How the Laws of Logic Lie

Gillian K. Russell

Australian Catholic University, Fitzroy, VIC 3065, Australia and University of St Andrews, St Andrews, Fife, Scotland, KY16 9AJ
Email: gill.k.russell@gmail.com

Abstract
Nancy Cartwright's 1983 book *How the Laws of Physics Lie* argued that theories of physics often make use of idealisations, and that as a result many of these theories were not true. The present paper looks at idealisation in logic and argues that, at least sometimes, the laws of logic fail to be true. That might be taken as a kind of skepticism, but I argue rather that idealisation is a legitimate tool in logic, just as in physics, and recognising this frees logicians up to use false laws where these are helpful.

> "We very soon got to six yards to the mile. Then we tried a hundred yards to the mile. And then came the grandest idea of all! We actually made a map of the country, on the scale of a mile to the mile!"
> "Have you used it much?" I enquired.
> "It has never been spread out, yet," said Mein Herr: "the farmers objected: they said it would cover the whole country, and shut out the sunlight! So we now use the country itself, as its own map, and I assure you it does nearly as well."
> Lewis Carroll, *Sylvie and Bruno Concluded* (1894)
> We should not be misled into thinking that the most realistic model will serve all purposes best.
> Nancy Cartwright, *How the Laws of Physics Lie* (1983)

I hope my title is legible, at least to some, as an homage to Cartwright's (1983) book *How the Laws of Physics Lie*, which argued that the use of idealisations in physics – things like point particles, frictionless planes, and perfectly isolated systems – led to many of its laws failing to be true. The present paper argues that logic makes use of idealisations too. Does that mean that there are false logical laws? Yes, it does – at least for some laws, and some logics. Out of context, saying that some of logic's laws are false can sound like a criticism of logic, but it is not intended as such here. Rather, I argue that idealisation is a legitimate tool and that recognition of its role can free logicians from the tyranny of truth, allowing

us to use idealisations – even idealisations that lead to falsehood – where these provide other benefits, such as simplicity, elegance, or tractability.

The paper can be understood as a contribution to a broader anti-exceptionalist project in the epistemology of logic that has been exploring the thesis that logic is continuous with the other sciences.[1] A central anti-exceptionalist theme has been that the epistemology of logic is, like that of other sciences, broadly abductive: logicians develop rival theories of their subject matter – the entailment relation – and assess them for accuracy as well as theoretical vices and virtues, such as simplicity or ad hoccery. Martin and Hjortland (2020) argued that logical theories, like scientific ones, are used for predication and explanation. These are central topics in the philosophy of science, but since Cartwright's book it has increasingly been recognised that so is idealisation. Shaffer (2012) takes one of her most important contributions to be 'her focus on idealisation as an important topic in the philosophy of science alongside more traditional topics like explanation, confirmation, realism, etc'. (p. 13) It is a natural next step for anti-exceptionalism then, to ask whether idealisation has a role to play in logic.

The paper is structured as follows: section 1 covers some preliminaries and sketches a way that a part of logic – model theory – can be construed as simulating the relation of logical consequence on a natural language. Sections 2 and 3 are the heart of the paper: 2 outlines and evaluates Cartwright's argument for thinking that the laws of physics lie and 3 looks at some ways idealisation is employed in first-order logic, and the laws which are false as a result. Part 4 responds to objections.

1. Preliminaries: simulating validity

Empirical scientists studying complex phenomena often build models; meteorologists build models of hurricanes and epidemiologists build models of epidemics. Models are used to predict future events – like windspeeds at landfall and infection rates at future dates – and explain past ones – like the hurricane's path and deathrates in a certain week. In this paper I will be talking about a particular part of logic – model theory – which uses the word *model* for a specific kind of set-theoretic structure. To reduce confusion, I won't also use that word for the things scientists build to understand hurricanes and epidemics. Instead I'll call these *simulations*.

Simulations of hurricanes and disease outbreaks are often computer simulations, which are most naturally understood as abstract programmes for turning a sequence of inputs: average wind speed, humidity, (ocean) surface temperature etc. into predictions. But scientific simulations can be physical as well: an orrery is a mechanical simulation of the solar system, and a river flow simulator a kind of angled sandbox with a water inlet at the top end, used to predict and explain erosion patterns and explore the consequences of interventions.

Computer simulations can be complex, but they are usually simpler than the phenomena they simulate. Air-temperature is a determinant of hurricane behaviour, but temperature varies between air packets and at the limit concerns the velocity of individual molecules. The *real* air temperature of a weather system is thus ferociously complex and a simulation would not attempt to represent all its detail. Instead, it will work with a value for average temperature over a certain packet of air. In this respect the simulation is simpler than what it simulates.

The relative simplicity is intentional and contributes to the usefulness of the simulation; we need to be able to run the simulation on inputs that we can collect or estimate

---

[1]See for example Hjortland (2017); Beall (2017); Russell (2019).

in a reasonable amount of time, and we need simulations to be computationally tractable, so that we get predictions in a timely fashion with available equipment. A simpler simulation that is easy to use is better than a more complex and accurate one that requires too much time and energy for the answers it provides.

Physical simulations are a little different. First, they *do* have extremely complex features that could in principle be measured or controlled. At the limit case, a physical model of a river could be an exact duplicate, down to the last detail. But such perfection is unachievable in practice, and often useless or even harmful.[2] Useful physical simulations are usually a different size from the target,[3] and we select a limited set of the target's properties to represent in the simulation. In the case of the river flow simulator perhaps the angle of descent, incoming water flow, or the shape of a curve. In an orrery it might be the relative size and orbital velocities of the planets.

Across history, different philosophers have thought of logic in a large number of different and incompatible ways: for example, as the normative study of reasoning, as the preconditions for thought as such, or as having no subject matter at all. Here I will follow Frege in *The Thought* in taking logic to be the science of truth.

> [Logic] has much the same relation to truth as physics has to weight or heat. To discover truths is the task of all sciences; it falls to logic to discern the laws of truth. The word "law" is used in two senses. When we speak of laws of morals or the state we mean regulations which ought to be obeyed but with which actual happenings are not always in conformity. Laws of nature are the generalization of natural occurrences with which the occurrences are always in accordance. It is rather in this sense that I speak of laws of truth. (Frege 1918: 289)

Physics studies the preservation or loss of heat over changes, logic the preservation or loss of truth over arguments. We can think of the model theory for a logic as simulating a complex phenomenon: the relation of validity on a natural language.[4] Natural languages are complex; verbs conjugate, nouns, and determiners decline, expressions exhibit vagueness, reference failure, self-reference, context-sensitivity, both lexical and scope ambiguity and sometimes natural languages are even thought to be self-contradictory.[5] The presence or absence of the logical properties – logical truth, logical consequence, logical equivalence, etc. – depends upon the syntax and semantics of the expressions which make up sentences and yet even in the simplest cases – names, *the*, *if*, *I*, *two* – these can

---

[2]Jorge Luis Borges' 1946 short story "Del rigor en la ciencia" ("On exactitude in science") concerns maps which are the same size as the territory they represent. "The following Generations, who were not so fond of the Study of Cartography as their Forebears had been, saw that that vast Map was Useless, and not without some Pitilessness was it, that they delivered it up to the Inclemencies of Sun and Winters." (Borges 1999) The idea is also treated in Lewis Carroll's final novel *Sylvie and Bruno Concluded* (1894), featured in this paper's epigraph.

[3]For example, an orrery is smaller than the solar system and a physical model of a cell is larger than the cell.

[4]Validity is an intuitive concept we have a pretheoretic grasp on, but developing a precise theory of it is a substantial task. Compare it, for example, to theorising effectiveness as a property of functions. On a standard story, effectiveness is an intuitive property that mathematicians tried to capture with various precise definitions, including Turing computability, recursiveness and abacus computability. Several of the precise theories turned out to define the same set of functions and this is sometimes taken to be evidence for the Church-Turing thesis: the thesis that the precisely defined set is the extension of our intuitive concept off effectiveness. But (the thought goes), such an identity is not amenable to proof, since one of its terms is imprecise and intuitive.

[5]E.g. Tarski (1944).

be difficult to theorise accurately. To study the logical properties then, model theorists build simpler languages with stipulated syntax and meanings: simulations.[6]

If we are looking for plausible examples of idealisation in logic there are various things we can point to. In tense logic we make simplifying assumptions about the way time is ordered or ignore special relativity. Indexical logics might assume that that every context of utterance has an audience. In the interests of concreteness and clarity, I am going to focus on a specific logic in this paper: an ordinary model theory for first-order logic (unary quantifiers but no functions or identity).[7] We begin with a specification of the language's simple ("primitive") expressions:

1. individual constants: a, b, c, etc.
2. individual variables: x, y, z, etc.
3. for each $n > 0$, non-logical predicates of arity n: P^n, Q^n, R^n, etc.
4. sentential connectives: $\neg$, $\wedge$, $\vee$, $\rightarrow$
5. quantifiers: $\forall$, $\exists$, …
6. punctuation: (,)[8]

Next we say something about forming complex expressions, here *well-formed formulas* (wffs):

1. If t_1, …, t_n are terms (variables or individual constants) and Π is an n-place predicate, then $\Pi t_1 \dots t_n$ is a wff.
2. If ϕ is a wff, then $\neg\phi$ is a wff.
3. If ϕ and ψ are wffs, then $(\phi \wedge \psi)$, $(\phi \vee \psi)$, and $(\phi \rightarrow \psi)$ are wffs.
4. If ϕ is a wff and ξ is a variable, then $\forall\xi\phi$ and $\exists\xi\phi$ are wffs.
5. Nothing else is a wff.

Helping ourselves to standard definitions of scope, variable-binding, and free-variable, we define a *sentence* as a wff with no free variables.

The next step is to simulate meanings on the expressions we have defined. This is done in at least three different ways. Some meanings are assigned by the interpretation function of a model. Let a model for this logic be an ordered pair

$$\langle D, \ I \rangle$$

---

[6]There is already a small literature on what Roy Cook calls "logic-as-modelling." Cook's own 2000 PhD dissertation advocates for the approach and he cites precursors for his work in Gentzen, Kreisel, and Lakatos, though it seems clear that both what Cook takes logic to be modelling (the behaviour and discourse of mathematicians) and the part of logic he takes to be the model (formalisation) is different from the version I present in the present paper. 'The idea, on the logic-as-model approach, is that formalization acts as a model of mathematical discourse, in much the same way that the Bohr model of the atomic is a model of the actual microscopic particles that make up the physical universe or a model ship is a model of a particular sea-going vessel'. (Cook 2000: 21) Shapiro (2001) also explores related ideas, though too he takes the target to be different from that assumed here: 'correct reasoning' which 'puts logic within epistemology' (rather than the laws of truth, as assumed here.) Still, Shapiro is clearly (as I am) thinking of an analogy with physics when he writes: 'I suggest that a formal model is a mathematical model of reasoning, in much the same sense, as a system of point masses is a model of moving objects. There are gaps between the model and what it is a model of'. (159)

[7]What follows here is intended to be quite standard; those familiar with these kinds of models will not lose anything by skimming this section.

[8]We allow individual constants and non-logical predicates to be subscripted so that we can always reach for a new one if we need it, and – for readability – to allow predicates to be written without superscripts and parentheses to be dropped where this does not cause other problems.

of which the first member is a non-empty set (the model's *domain*) and the second a function from the non-logical expressions to suitable extensions on that domain (the model's interpretation function.)

I is a function such that:

1. If t is an individual constant, then $I(t) \in D$.
2. If Π is an n-place non-logical predicate, then $I(\Pi) \subseteq D^n$.

Think of the interpretation function as assigning (idealised) meanings to the non-logical expressions. They are idealised in that, in the simulation, for example, individual constants never lack referents, or have more than one.[9] Similarly the extension of an *n*-place predicate is always a determinate subset of D^n (we will see a contrast case for this claim in §3.)

But not every expression in the simulation gets its meaning from an *I*-function. Variables get their extensions from a *variable assignment*: a function from variables to elements of the domain. The meaning of a variable is thus dependent on, but not completely determined by, the model. If we write $[t]^g_M$ to mean the denotation of the term, t, in a model, M, on a variable assignment, g, then:

$$[t]^g_M = \begin{cases} I(t) & \text{where } t \text{ is an individual constant} \\ g(t) & \text{where } t \text{ is a variable.} \end{cases}$$

And finally, crucially, the logical expressions get their meanings independently of models: they are assigned via clauses in the recursive definition of truth-in-a-model-on-an-assignment (we write $V^g_M(\phi) = 1$) which apply across all models:

1. $V^g_M(\Pi t_1 \dots t_n) = 1$ iff $\langle [t_1], \dots, [t_n] \rangle \in I(\Pi)$
    a. $V^g_M(\neg\phi) = 1$ iff $V^g_M(\phi) = 0$
    b. $V^g_M(\phi \wedge \psi) = 1$ iff $V^g_M(\phi) = 1$ and $V^g_M(\psi) = 1$
    c. $V^g_M(\phi \vee \psi) = 1$ iff $V^g_M(\phi) = 1$ or $V^g_M(\psi) = 1$ (or both)
    d. $\vdots$
    e. $V^g_M(\forall\xi\phi) = 1$ iff for all $o \in D$, $V^{g^\xi_o}_M(\phi) = 1$.[10]
    f. $V^g_M(\exists\xi\phi) = 1$ iff there is $o \in D$ such that $V^{g^\xi_o}_M(\phi) = 1$.

The meanings of the logical constants are also plausibly idealised; perhaps the meanings of the natural language *and*, *or*, *if*, and *all* are more complicated than this.[11] We might even think that the base clause in the definition introduces an idealisation: we assume that an atomic sentence can only have one of two truth-values, 1 and 0, and make no

---

[9]Whether or not something qualifies as an idealisation depends on what the target phenomenon is like. If the physical universe contained perfectly spherical planets then the assumption that a planet is a perfect sphere in a simulation would not thereby be an idealisation. Similarly, if every natural language name had exactly one referent, then names in the formal language would not be idealisations in virtue of having that property themselves. There is thus room for disagreement about the *extent* to which different formal languages idealise natural language. Since my goal here is to understand – through illustration – how idealisation can work in logic, I will sometimes assume natural but admittedly controversial positions about natural language meanings without argument.

[10]g^ξ_o is the variable assignment just like g except that $g^\xi_o(\xi) = o$, a so-called ξ-variant of g.

[11]Though see §3 for more on how this relates to Grice's views at the beginning of 'logic and conversation'.

distinction between say, a sentence which expresses a false proposition and one which fails to express any proposition at all. Finally, those who think that some sentences can be both truth and false will also think that this simulation omits a complexity which is present in the case of natural language sentences.

The above allows us to define truth in a model:

Truth in a model
A sentence ϕ is *true in a model M* just in case $V_M^g(\phi) = 1$ for all assignments g. We write $V_M(\phi) = 1$.

And now we can define the logical properties:

Logical truth and logical consequence
A sentence ϕ is a *logical truth* if it is true in all models. We write: $\vDash \phi$
A sentence ϕ is a *logical consequence* of a set of sentences Γ if all models of Γ are models of ϕ. We write: $\Gamma \vDash \phi$.

I think it plausible that the model theory above simulates features of natural language, much as a computer simulation representing processes of convection, radiation, and conduction in a room can simulates what happens to the temperature when a fire is lit. But I do not want to maintain that every kind of model theory in logic is best understood as a simulation. Some may be so abstract, or so unlike natural language meanings, that they are better thought of as black boxes which take formal sentences and arguments as inputs and yield verdicts on their validity. Here I focus on a standard, Tarski-inspired model theory for FOL precisely *because* it is so naturally understood as attempting to simulate the relation of logical consequence on a natural language.

I need to address two final preliminary issues before moving on to look at Cartwright's argument in the next section: (i) what the entire model represents, and (ii) what a *law* of logic is.

In the theory above, a model is a pair: $\langle D, I \rangle$. It's well known from Etchemendy (1990) that there are two standard views about what such models represent: on the metaphysical approach, models represent different ways the world could be – different *possible worlds*. On the semantic approach, they represent reinterpretations of the non-logical primitives – what we might call different *possible languages*. These correspond to two rival conceptions of what logical truth is: on the metaphysical view it is truth in all possible worlds – necessary truth – and on the semantic one truth in every language. I have argued elsewhere that neither view is correct: instead models represent *combinations* of worlds and languages, so that logical truth is (roughly) truth in all worlds *on* all languages. (Russell 2023; Shapiro 2009; Sher 1996) Here I am trying to focus on different issues, so I will simply note first, that there are competing views about what a model represents and second, that on each of these models are clearly highly idealised: a pair $\langle D, I \rangle$ makes for a very austere representation of a possible world, or a possible language, or a combination of the two.

And so finally, to the *laws* of logic. The definitions of logical truth and consequence quantify over models: e.g. $\vDash \phi$ iff ϕ is true in *all* models, and $\Gamma \vDash \phi$ just in case *all* models that make Γ true make ϕ true. The presence of these properties thus depends on more than what happens in the 'actual' model (if there is one.) Something analogous happens with scientific simulations too. When it comes to establishing laws of physics, such as the ideal gas law or the GTR field equations, these are meant to hold over all the simulations, not just in the one that most closely represents the actual world.

Though there is not much regimentation in usage here, simple principles in which $\vDash$ is the main predicate, like $\vDash \phi \vee \neg\phi$ (LEM) and $\phi \rightarrow \psi$, $\phi \vDash \psi$ (MP), are quite naturally dubbed *laws of logic*[12] when there are no countermodels. This is how I will use the phrase in this paper. The laws of logic according to some model theory then, are the principles of the form $\Gamma \vDash \phi$ (Γ may be empty). These do not hold *relative* to a model, but rather count as laws because they hold throughout the class of models: i.e. every model in which Γ is true is one in which ϕ is too.

2. How the laws of physics lie

Cartwright argues that some laws of physics lie, but not all of them. She distinguishes *phenomenological* from *theoretical* laws, where the former 'describe what happens' in specific physical situations e.g. 'what happens in superfluids or meson nucleon scattering' (Cartwright 1983: 2). If what they describe happening is what actually happens, then phenomenological laws are true.[13]

Theoretical laws, meanwhile, are the abstract equations of fundamental theories, which seek to explain what happens, not merely describe it. Examples include the equation of continuity and Boltzmann's equation which 'are thoroughly abstract formulae which describe no particular circumstances'. (11) Cartwright argues that these laws are not true, since they are supposed to explain and 'paradoxically enough the cost of explanatory power is descriptive adequacy. Really powerful explanatory laws of the sort found in theoretical physics do not state the truth'. (3)

She provides several interrelated arguments for this thesis. The most central is to do with the difficulty of directly testing theoretical laws. These laws have no direct consequences for experiment on their own, and so need to be connected to experiment with bridge principles, which tell us what we would expect to find if the laws were true. The bridge principles are various and complicated and must be applied intelligently. They are not given by the fundamental law itself, and so the law itself does not tell us what the results of experiments will be – it does not 'describe the world" and so isn't true. On Cartwright's view the theoretical laws are more about finding a good way to organise our view of the world than about matching what is out there.

But the idea of Cartwright's that I want to focus on receives somewhat less attention in her book. It is the idea that the theoretical laws in physics are not true *because they use idealisations*. As she puts it in one place: 'The phenomenological laws are indeed true of the objects in reality – or might be; but the fundamental laws are true only of objects in the model'. (4)[14]

This thought is highly suggestive, but – to a philosopher of logic and language at least – non-trivial to spell out in satisfying detail. On one way of thinking about it, geometry also deals with idealisations. It talks of straight lines and triangles, which are never found in the physical world: look closely enough at any physical edge and it will be revealed to be at least a little curved, or a little bumpy. If the edge is part of a closed 3-sided figure, that figure can only approximate a true triangle. We might

---

[12]E.g. *Law* of excluded middle. See also Russell (2018: §1.2–2).

[13]Despite what the name, Cartwright's use of *phenomenological* does not limit it to things that can be experienced or even the observable, but rather to what can be established with experiment: "Observation – seeing with the naked eye – is not the test of existence here. Experiment is". (6)

[14]This idea is endorsed elsewhere, see e.g. (Shaffer 2012: 2) 'If theoretical claims hold true only in idealised models, then such theoretical claims are not even approximately true. Therefore, theoretical claims are not even approximately true'.

be tempted to say that an equation of geometry, like the Pythagorean formula – which tells us that $a^2 + b^2 = c^2$ (that the square of a triangle's hypotenuse is equal to the sum of the squares of the other two sides) is 'only true of objects in the model'. But – here's the problem – the formula is a (somewhat disguised) universal generalisation. If we make that explicit it begins "*for all triangles*, the square of the hypotenuse is equal to the sum of the squares of the other two sides." And if triangles are idealisations, and there are no triangles in the physical world, then the universal claim is true of the physical world – just vacuously so. That suggests that the use of idealisation leads to the laws being vacuously true. And vacuously true is not false.

Similarly in physics. The ideal gas law is an equation relating the temperature, pressure, and volume of an ideal gas:

$$pV = nRT$$

where p is the pressure, V is the volume, T is the temperature, n is the amount of substance (in moles), and R is the ideal gas constant. What is an ideal gas? It is a bunch of point particles moving around and never interacting with each other, in a closed system. This incorporates at least three different idealisations: the point particles themselves, the assumption that they never interact, and the assumption that the system is closed (no real container of gas could be completely isolated from outside forces.) A plausible view is that although some physical gases approximate ideal ones in their behaviour, there are no ideal gases in reality. The ideal gas law is – like the Pythagorean theorem – an implicit universal claim. It says that all ideal gases are such that $pV = nRT$, and so, given that there are no ideal gases, it is vacuously true. If a law is vacuously true, it might be problematic in other ways, but it does not *lie*.

Still, it is clear reading Cartwright that she thinks that theoretical laws lie because they are *not true*. She would not be satisfied with *vacuously true*. My sense from reading her book is that she is a bit looser with 'true' than philosophers of logic and language like to be, and sometimes uses it as if it is synonymous with *describes reality*. We might well say that vacuously true laws do not 'describe reality' and so this interpretation of what Cartwright has in mind would at least explain why *she* says that the laws are not true.[15,16]

But if we do not want to adopt that view of truth (and I do not), then so far we have an argument for thinking that idealisation can lead to laws which are true in an idealised simulation, and for thinking that they are only vacuously true of the physical world. This sounds bad, but it falls short of an argument that the laws are not true. Can we do better?

A claim can fail to be true in two ways: by failing to have a truth-value at all – perhaps because of failure of meaningfulness – or by being false. I will consider each in turn. If the laws employ empty names (perhaps *Vulcan* is one) we might be able

---

[15]Of course, if this were sufficient for the laws to lie, then those who hold that logic has no subject matter and is not about anything – reality included – will hold that the laws of logic do not describe reality and so are not true. I think this just means that using *true* as a proxy for *describes reality* might be something you can get away with in the philosophy of physics, but which will lead to trouble in the philosophy of logic.

[16]The issue is complicated by the fact that she also says – in contrast to other philosophers of science like Van Fraassen – that she is happy to say that theoretical entities *exist*. 'One can reject physical laws without rejecting theoretical entities'. (6) But the theoretical entitles that Cartwright believes in are not point particles and isolated systems, but e.g. molecules and tangential stresses; theoretical entities that we cannot see directly but which experiment reveals to us as causes of what happens. Molecules are theoretical entities without being idealisations. Point particles are idealisations and, like triangles, they do not exist in the physical world.

to make a case that they lack a truth-value on the grounds that they fail to express a proposition (e.g. perhaps *Vulcan is 5000 km in diameter* has no truth value because fails to make a claim about anything.) Laws with no truth value are not true.

However, the generalisations we have been looking at are not like this: *triangle* and *ideal gas* appear to function as common nouns, which are usually treated as predicates. Sometimes people suggest that a predicate will be meaningless if there is nothing in the world to which it applies – one sometimes hears this from Wittgensteinians and ordinary language philosophers who think that meaning is use and *use* means used for things we really encounter.[17] But I think it deeply implausible that *triangle* is meaningless, even though speakers do not encounter triangles. Rather, they become acquainted with imperfect physical triangles (perhaps the faces of wooden blocks or shapes drawn in crayon) and gradually come to understand that a triangle is *like that* but with perfectly straight edges. We do not see examples of perfectly straight edges either, but we come to understand the requirements by learning the ways in which perfect straightness can fail – being slightly curved or bumpy – and grasping the complex condition: like that (the imperfect line) but with no curves or bumps. Go on taking bumps and curves away, and at the end of that road is *straight*.[18] So the prospects for arguing that the laws of physics are *meaningless* because of idealisation look slim.

What about arguing that they are false? Since the laws are universal generalisations – of the form *All Fs are G* – they would be false if there were Fs which were not G. It often feels as if it is being suggested that it is something to do with the fact that *real gases* do not obey the ideal gas law that makes the ideal gas law false. But if the law is only *about* ideal gases, the deviant behaviour of non-ideal gases is no obstacle to its truth. So what if it is *not* only about ideal gases? It does seem odd for physics to rest easy with laws governing ideal objects and deny making claims about the physical world. So suppose instead that the ideal gas law should be understood as saying:

$$\text{for any gas,} \quad pV = nRT.$$

This *is* false. So the physicist faces a dilemma: either the law quantifies over ideal gases alone – in which case it does not talk about physical reality and can hardly be called a law of physics – or it quantifies over real gases and is false. So the laws *of physics* lie.

This, I think, is the strongest construal of the argument that idealisation in physics leads to laws that lie. The laws in question take things that are only true of idealisations, and *say* them about things in the physical world. It will always be a temptation to defend such laws by restricting their application to the idealisations – to say that they are true 'of the model'. But that retreat is a tempting mistake: it relinquishes the claim to have been giving laws *of physics* at all.

Above I have usually spoken as if an idealisation is a special kind of object or property – like a point particle – but there is a different way of thinking of them, as assumptions. In her recent book on idealisation in science, Potochnik writes: 'Idealizations are assumptions made without regard for whether they are true and often with full knowledge that they are false'. (Potochnik 2017) Natural examples include the conventional assumption in STR that the speed of light is the same travelling away from an observer as it is returning, and the assumption that a local physical system is isolated from

---

[17]Norman Malcolm, for example, suggests that it must be commonplace for people to be certain about empirical statements because otherwise people would be unable to learn what *certain* means.
[18]Russell (2022: 816–17).

outside forces. I suspect the boundary between idealisations-as-objects/properties and idealisations-as-assumptions is not very robust. The point particles in an ideal gas do not interact; we can describe those particles as a special kind of object, or instead think of ourselves as making an assumption: point particles do not interact. Perhaps every idealisation can be described either way. But conceiving of idealisations as assumptions does make it easier to see how they result in falsehood. For first, if the assumption is part of the theory and false, the theory is false. That way of thinking of a theory makes it seem like a set of sentences, counted as false because one of its members is. An alternative is to think of a physical theory as a set of simulations – representing the set of possible evolutions of the physical system – and a law as an equation (or other principle) that holds in all of the simulations. Roughly, the more assumptions we make, the fewer simulations we include, and the larger the set of equations which come out true-in-all-simulations and hence as laws. Where our assumptions are false, we will tend to exclude models we should not, and hence have more laws than we should: some of the laws of will lie: they are false in situations that we are ignoring for convenience or in the full knowledge that they are possibilities.

3. Idealisation in the model theory for FOL

Model theory differs from physics in ways that might suggest that Cartwright's overall picture could find no purchase there. There are no causal or phenomenological laws in logic and truth is not concrete, physical, or measurable in the way force, temperature, and charge are. I do not deny that there are differences; even the most ardent anti-exceptionalist should not think that logic *is* physics. Still, logic is similar to physics in two respects: (i) we use idealisation in both logic and physics, and (ii) in both this sometimes results in laws which are false; they are laws according to the theory (none of the theory's models are counterexamples) but *really* they are false. In this section I explore four examples of idealisation in the FOL model theory from §1 and show how each leads to false laws.

3.1. Idealization 1: non-empty domains

In the theory from §1, a model is a pair in which the first element is a non-empty set of elements: the model's domain. The requirement of non-emptiness is an idealisation. It excludes models which represent circumstances where nothing exists. This results in a much simpler theory. But it also gives us the following logical laws:

1. $\forall x Fx \vDash \exists x Fx$
2. $\forall x Fx \vDash Fa$
3. $\vDash \exists x(Fx \rightarrow \forall y Fy)$
4. $\vDash \exists x Fx \vee \exists x \neg Fx$

1. and 2. have no FOL-countermodels because, in each case, every FOL-model of the sentence on the left of the turnstile is a model of the sentence on the right. 3. and 4. have no FOL-countermodels because each sentence to the right of a turnstile is true in all FOL-models.[19] Still, I submit, this is only because FOL excludes models with empty domains. If the domain could be empty, $\forall x Fx$ would be trivially true, but $\exists x Fx$ false. Such a model would be a counterexample to 1. And also to 2, 3, and 4.

---

[19]The third is a well-known 'trick' case. It works in part because it existentially quantifies into a conditional, which results in a very weak sentence.

A defender of these laws may be tempted by a response here: but there *are* no FOL-models with empty domains! We defined our models so as to exclude them. This I think is the counterpart in logic of Cartwright's observation that the theoretical equations are true only 'in the models'. When we find situations not represented by the official models where the equations fail, we retreat to saying that the laws only ever made claims about the models, and so are not challenged by such examples. But in physics this retreat made the laws no longer qualify as laws of physics; laws which do not attempt to describe the physical world do not qualify.

The situation is similar in logic. The retreat to speaking only of relative-validity, i.e. valid-in-FOL, backs off the challenge of theorizing the target phenomenon, validity. The relative properties (valid-in-FOL, valid-in-intuitionistic logic, valid-in-LP etc.) are of interest, but they are not the target: they are each attempts to theorize the target. If an argument does not preserve truth in some circumstances, i.e. if there is a way to make the premises true without making the conclusion true – e.g. circumstances where nothing exists – then it isn't *really* valid.[20]

So the non-empty domains assumption is an idealisation, and it is one that leads to laws of logic that lie, namely, 1, 2, 3, and 4 above.

3.2. Idealisation 2: all terms have a denotation

The representation of names in FOL is also idealised, in that each name is assumed to have a unique denotation. Natural languages have empty names, e.g. names which fail to denote an object, like *Vulcan*, or *Homer*. But in FOL names are simulated by the non-logical individual constants, *a*, *b*, *c*, … etc. and each is assigned a unique element of the model's domain by the *I*-function. This is a helpful simplification. It saves the need to consider what truth-value a sentence ought to receive if a term it contains fails to denote. Some might think such a sentence should be false, or that it should lack a truth value. FOL sidesteps these issues, but one of the consequences is that certain laws hold that might not otherwise have held. For example:

5. $\forall x Fx \vDash Fa$
6. $Fa \lor \neg Fa$

3.3. Idealisation 3: predicates are complete

A different idealisation concerns the relationship between the extensions of predicates and the extensions of their negations. In FOL the extension of a primitive non-logical, unary predicate, Π, is a subset of the domain,

$$I(\Pi)$$

and the extension of the negation of that predicate (the predicate's *anti-extension*) is

$$D - I(\Pi).$$

---

[20] Again I think the comparison with effectiveness can be helpful: if we proposed the class of primitive recursive functions as the extension of *effective*, we might be challenged by examples of intuitively effective functions which are not primitive recursive, e.g. partial recursive functions. An imaginary interlocutor might maintain that they had intended to talk only of 'effective according to my theory on which the effective functions are the primitive recursive ones' but with this move they back of the challenge of theorising the target.

The extension and anti-extension thus exhaust the domain: every element is either in the extension or the anti-extension.

$$I(\Pi) \cup (D - I(\Pi)) = D$$

It is not clear that natural language predicates interact with negation like this, for various reasons. Some philosophers think that vague predicates may be *incomplete*, i.e. there may be objects which are neither in the extension nor in the anti-extension. Take colour terms. It might be that there are some shades in the extension of *red*, e.g. scarlet, and some in the anti-extension of *red*, say the brightest orange, but that some shades in the middle of the spectrum between scarlet and orange are neither in the extension of *red* nor of *not red*. As we might think of it: the linguistic rules don't say enough to assign the shade to either camp: *red* or *not red*. (Soames 1999: 167)

Perhaps something similar happens with category mistakes. A category mistake occurs when we try to apply a predicate to an element from outside of the category for which the predicate is defined. For example, when we ask whether zero is orange, or whether the interrobang (‽) comes at the beginning or end of the alphabet. Zero does not satisfy the predicate *orange*. But, one might think, it does not satisfy *not orange* either – 'orange' is not defined on abstracta, and the interrobang is not the kind of thing that gets a position in the alphabet.

A third home for incomplete predicates is in theories of truth, where we find a motivation which is apparently independent of vagueness or category mistakes. Kripke's 'Outline of a Theory of Truth' used incomplete predicates to simulate the meaning of *is true*. He writes:

A sentence such as (1) ["Most (i.e., a majority) of Nixon's assertions about Watergate are false"] is always meaningful, but under various circumstances it may not "make a statement" or "express a proposition." (I am not attempting to be philosophically completely precise here.)

To carry out these ideas, we need a semantical scheme to handle predicates that may be only partially defined. Given a nonempty domain D, a monadic predicate $P(x)$ is interpreted by a pair (S_1, S_2) of disjoint subsets of D. S_1 is the extension of $P(x)$ and S_2 is its anti-extension. $P(x)$ is to be true of the objects in S_1, false of those in S_2, undefined otherwise. (Kripke 1984: 699–700)

Kripke's paper shows how to construct a hierarchy of languages, each with a truth predicate that extends that of the last, until we reach fixed points, at which languages contain their own truth-predicates. Yet even at the fixed points, the truth-predicate – thought intuitively 'more complete' – is not 'completely complete': ungrounded sentences, including the Truth-Teller and the Liar, are neither in the extension nor the anti-extension of *true*. (Kripke 1984: 707)

It is fairly natural to think of the completeness of a predicate as coming in degrees, much as the straightness of a line does, or the flatness of a surface.[21] Truth-predicates earlier in Kripkean hierarchies of languages are *less complete* than those that come later

---

[21]*Complete* is an absolute expression – like *flat* or *straight* – which means that satisfying it is a matter of being at the limit on some associated scale. Following the standard pattern for such expressions, *more*

(i.e. the degree of completeness is less.) FOL deals only with predicates at the limit on this scale: fully complete predicates for which the union of the extension and the anti-extension exhausts the domain. On the assumption that natural language exhibits the features canvassed above, FOL predicates are idealisations – like straight lines and flat surfaces.

The most obvious law arising from this idealisation is the law of excluded middle.

7. $\vDash Fa \vee \neg Fa$

where F is an incomplete predicate and a denotes an element neither in the extension nor the anti-extension of F, this FOL law is not true, since neither Fa nor $\neg Fa$ is true.

3.4 Idealisation 4: truth-functional connectives

The language laid out in §1 contained $\neg$, $\wedge$, $\vee$, and $\rightarrow$ as its only sentential connectives. Each received its meaning from the recursive definition of truth-in-a-model-on-an-assignment, which was a less pictographic way of assigning them the truth-functions depicted by the familiar truth-tables:

ϕ	$\neg\phi$
T	F
F	T

ϕ	ψ	$\phi \rightarrow \psi$
T	T	T
T	F	F
F	T	T
F	F	T

Grice's transformative essay 'Logic and Conversation' taught philosophers to be cautious about attributing complicated meanings to the natural language expressions based on complicated usage patterns.[22] In particular, he argued that it was naive to think *either* that the formal connectives were a scientific improvement on their natural language counterparts, *or* that they were a useless over-simplification because in fact $\vee$ and *or* have the same meaning. The difference in their use is to be explained by implicature, rather than truth-conditions.[23] If Grice is right, then the meaning of $\vee$ in FOL is not an dealisation of the meaning of *or* in English.

I accept the point that difference in use is compatible with sameness of meaning, and so I shan't argue *on the basis of use* that the connectives in FOL are idealisations of their natural language counterparts. This means that the key to arguing that two expressions have different meanings lies elsewhere. For example, we might show that substituting one for the other in a sentence can take that sentence from true to false, or that the

---

complete is roughly synonymous with *closer to being complete* and *x is less complete than y* entails *x is not complete* (Kennedy and McNally 2005; Unger 1971).

[22]Grice (1991). Originally given as a lecture in 1967 then published in 1975.

[23]For example: we do not normally transition between, say *Issa went to Zanzibar* and *Issa went to Zanzibar or Issa went to Arusha* in ordinary conversations because the latter is less *informative* than the former, and so violates Grice's maxim of quantity. See (Soames 2003: Ch. X) for more on 'Logic and Conversation' and its implications.

epistemic and semantic consequences we expect when two expressions have the same content do not obtain.

In the previous section already we saw some reasons to think that the FOL treatment of negation is simplified: if some predicates are incomplete, then *Fa* could fail to be true without ¬*Fa* being true. The FOL truth-table for negation leaves no space for this: for any sentence ϕ, if ϕ is not true, then $\neg\phi$ is and vice versa. What we need is a third truth-status – call it *I* (for *indeterminate*) – and a row of the truth table that tells us what happens if ϕ has that status. Kripke used Kleene's 'strong' tables, which give us:

ϕ	$\neg\phi$
T	F
F	T
I	I

The classical table is clearly a simplification of this – it is what we get if we add the assumption 'every sentence is true or false' to our theory and thereby excise a set of models – the set of models relative to which some sentence is assigned *I*. So we were already working with an idealisation of *not* in the previous section, and once we have added a third truth-status, the FOL tables for the other connectives specify only partial functions: we need to say what truth-status each yields when one or more of its arguments is indeterminate.

Thanks to strong Kleene logic (K3), we have a good idea what happens to the laws when we take the new cases into account. All laws that had non-empty premise sets survive; all logical truths die. So – in addition to 7. above – the laws that lie thanks to the idealisation of having only two truth-statuses include:

8. $\vDash \neg(\phi \wedge \neg\phi)$
9. $\vDash \phi \leftrightarrow \neg\neg\phi$
10. $\vDash \phi \leftrightarrow \phi$

Each comes out indeterminate on an interpretation that assigns *I* to each atomic sentence contained in the principle.

A second promising fault-line for arguing that some FOL-connective is an idealisation is the conditional. In linguistics, Kratzer's modal account of natural language conditionals has been dominant for many years now, and this is both a somewhat more complicated account than the material one, and one for which the material conditional seems like a sensible simplification for certain purposes.

4. Objections

A central goal of this paper is to sketch a picture of what it is we are doing when we do model theory. It is a picture on which certain troubling problems disappear. For example, it can seem distressing to relinquish the elegance and strength of classical logic for the complexities of a non-classical system – even when the very strength of the classical system seems to lead us into paradox. But if logics simulate distributions of truth over natural language sentences – he way meteorologists simulate wind-speeds over hurricanes – we can help ourselves to Cartwright's insight that "[w]e should not

be misled into thinking that the most realistic model will serve all purposes best." (152) In particular, we can allow that classical logic is not the most realistic model, that it gets some things wrong, and that some of its laws are false, *without* the distress of abandoning it, since it may still be the appropriate model to use in many circumstances, thanks to its simplicity, tractability etc.

One way to object to this picture is to target the assumptions that it makes about logic. Hurricanes, gases, and disease outbreaks are real, physical things, out there in the world. My picture of logic might seem to suggest then that logic is out there, waiting to be discovered, like a disease. But there are extant views in the philosophy of logic on which it is not, on which logic is a human invention, a mere convention, or perhaps a mental construction, as some philosophers take numbers to be. And so we can ask: does the view sketched here presuppose realism about logic, and is that legitimate?

It is complicated. Thanks to the inspiration from empirical science, the simulation and idealisation view presented here is a natural fit with a pretty robust realism about consequence and, beyond this, the very idea of an idealisation is relative to a real thing represented more or less accurately. If all gases *were* ideal, ideal gasses would not be idealisations after all. Similarly, if all predicates were complete, then FOL's complete predicates would not be an idealisation, and so on. So there being something real which is represented by the simulation is a presupposition of the picture presented here.

But there are at least three kinds of thing we might be realists (nor not) about: logical consequence itself, truth, and the other things represented by the simulation, using e.g. elements of a domain, interpretation functions, clauses in the definition of truth-in-a-model, predicates, individual constants, variables, parentheses, and – thinking beyond FOL – the possible worlds, times, accessibility relations, and multiple domains of more complex models.

The overall simulation and idealisation picture is clearly compatible with various views about many different parts of the simulation. Perhaps variables are not an idealisation of something in the world, just a useful instrument for stating an idealised meaning for quantifiers. We could be instrumentalists about variables. Similarly, perhaps we do not know whether anything corresponds to the 'possible worlds' (elements of a domain W) in a Kripke model, and so for now we use these instrumentally to capture patterns over the truth-values of sentences containing certain modals.

Truth, I think, has a special status, as the property whose laws logic tries to uncover. To hold the view I have sketched without thinking that truth is real seems quite odd to me. That would be like being a physicist who wants to uncover the laws of motion but thinks there is no such thing, or the laws of heat preservation while they are a skeptic about heat. Perhaps there is space for different views about what heat ultimately is (the presence of a special substance, vs mean molecular motion) but the project of trying to discover how it behaves presupposes the reality of the thing being studied.

The truth-property in question holds of sentences, and sentences are human constructions, but I do not think them thereby less real than hurricanes or diseases. Linguists also study natural languages, and they too find various ways to simulate them. Being a property of a human construction is no bar to being studied through simulation and idealisation.

The question remains whether logical consequence itself is real, and this perhaps depends on the resolution of particular controversies concerning what it is – a complex modal property, a semantic property, or a combination of both – as well as the metaphysics of the modal and the semantic. Some philosophers (e.g. Quine) are skeptical or irrealist about the modal and the semantic (and personally I incline to a sort of

optimistic realism about both.) But I think there is room in the simulation and ideal-isation picture for some ecumenicism here. If we think heat is real, we can simulate dif-ferent possible states of the world and extract the laws of heat from our models, and still disagree about the status of *the laws*. It's the same with the laws of truth. So to summar-ise my response to this kind of objection: the view presupposes some kinds of realism, fits well with others, and requires no verdict on still others. I am quite content with my commitment to the presupposed kinds: realism about sentential truth, and realism about the targets of idealisations – things like predicates and referents.

A different objection is happy to concede that truth and what we might call *analytic* (or sometimes *semantic*) consequence are real phenomena and suitable targets for simulation, but suggests that *logical* consequence is something more refined, something forged in the fires of mathematical logic, rather than encountered in the everyday world. Though we *can* simulate natural languages and their consequence relations, that is not what logicians are doing with, e.g. first-order classical model theory. In support of this objection we might note that mathematical logicians do not usually think of themselves as concerned with natural languages; their discipline is largely autonomous.

There are really two related issues here – both of which could support book-length treatises, and both of which can potentially support objections to my view: there is the question of whether logical and analytic consequence are independent targets (with my view perhaps only suited to the latter), and the (potentially separable) idea that math-ematical logic is autonomous. Let me start with the former.

A standard way to draw a distinction between logical and analytic truth is with refer-ence to the *logical constants*: a sentence is an analytic truth if it is true in virtue of the meanings of the expressions it contains, but a logical truth if it is true in virtue of the meanings of a subset of the expressions it contains – the ones which are logical constants. (This story connects neatly to the different ways we treat expressions in model theory: logical expressions keep their meanings over models, non-logical ones do not. If a sen-tence remains true over all models then its truth cannot have been dependent on the meanings of the non-logical expressions, only on the meanings of the logical ones.) Standard natural language logical constants include *and*, *if*, *every* and *not*, but sometimes more are used: *must*, *knows*, and even *is a member of* and indexicals like *I*, and *now*. (Their formal counterparts: $\wedge$, $\rightarrow$, $\forall$, $\neg$, $\Box$, K, $\in$, i and N respectively.) Analytic and logical truth are clearly closely related to ideas, and distinguishing them turns on explain-ing what the logical constants are. That is, not just giving a list of expressions which we will treat as logical constants (as we do when we introduce a model theory) but an account of what features make an expression such that it ought to be treated as a logical constant, and such that the resulting subcategory of analytic truths – those true in virtue of the meanings of the logical expressions – is interesting enough to be worthy of inde-pendent study. Such an account has proved difficult to come by. MacFarlane (2009) pro-vides a helpful overview of the failures of the leading approaches. See Russell (2023: Ch. 10) for my own view. Tarski (1936) himself, with endorsement from Varzi (2002), sug-gested that in principle any expression could be treated as a logical constant,[24] making our choice of constants in any one case a pragmatic, conventional matter, and the dividing line between logical and analytic consequence conventional as well. That is, on such a view, there is no *real* distinction between logical and analytic truth, just the practical ques-tion of where we ought to draw the line (i.e. which expressions we ought to treat as logical.) While there is a great more deal to be said about this issue, I hope I've said

---

[24]Tarski also provides a different account in Tarski (1986).

enough to suggest that it's far from clear that there is a persuasive objection here: even if there is a real distinction to be drawn between expressions which ought to be treated as logical, and those which ought not, and thus a distinction between logical and analytic truth that is more than pragmatic, it is far from clear that logical truth is such a different beast that it is impossible to study it through the present lenses of simulation and ideal-isation, and so gain the benefits of so-doing.

That leaves the separable issue of the autonomy of mathematical logic from the study of natural language consequence. It is a *highly* autonomous discipline. Expert research-ers in formal logic – even model theorists – needn't have much interest in natural lan-guage phenomena and often investigate questions simply because they are interesting, with no thought for whether this helps us understand natural language consequence. This seems different from many meteorologists, whose interest in simulations of hurri-canes is supposed to be motivated and justified by their interest in real hurricanes.

Meteorologists however, *can* become interested in the simulations themselves. They might investigate the mathematics or computer science of their simulations, including e.g the relative complexities of the simulations or the processing power required for certain answers, or they might study risk factors for developing category four status *by studying models alone*, without ever 'stepping out' of the simulations to compare their results with the real world. It is, I think, pretty standard academic behaviour to become interested in questions about and within the simulations themselves. (Some people might think this is getting distracted, others that we are finally getting started.) This makes for a very natural just-so story about the origins of geometry: we began with an interest in the practical and physical: how much seed it takes to sow a field, whether there is a fair way to split a loaf of bread, which of two containers holds more apples, how much twine we need to wrap all the bundles. If we make good progress we might end up with a general theory like Euclidian geometry, which talks of points, lines, and planes – denizens of a simulation – and which can be investigated further, quite autonomously, without consideration of fields, or bread, or twine. But Euclidean geometry has rivals – such as Riemannian geometry – and if we are interested in the question of which is true, then is time to go back to the physical measuring devices – meter rules, light-rays, and clocks – to adjudicate. Similarly where there are rival models of entailment – different logics. Questions about the models can be independently interesting – interesting enough to ground a whole research area – but the autonomy of formal logic is not an argument *against* the simulation and idealisation view: it's simply a consequence of the fact that the simulations are independently interesting.

This sets up us up for the last objection that I will consider here. In trying to precisify Cartwright's argument, we asked whether the laws were about the things we are simu-lating, or only about the simulations. For example, does the implicit universal quantifier in the ideal gas law quantify over all gases or only over all ideal gases (of which there are none). I suggested that we can see this as setting up a dilemma: if it quantifies over all gases, the law is false, but if it only quantifies over things that do not exist (ideal gases) then it is vacuously true, as well as not really about the physical world after all. Thus to the extent that the laws are laws *of physics* they lie.

Here is the objection: there is a third option. The quantifier is neither about physical gases, nor vacuous, but about ideal gases – abstract objects 'in the model'. This fits with Cartwright's own tendency to put it in terms of the laws of physics not being true, but only being 'true in the model'. (Something that she clearly does not take to entail that they are *true*.) It is also quite a plausible thing to say about elementary geometry: it is about points and lines and planes – abstract things that might not be physically instan-tiated. So perhaps we should say something like this about the law of excluded middle,

i.e. it's not *about* how things have to be in natural language, but only about the simulated languages with their restricted interpretations and idealised models. Of course, there are different simulations, which include different classes of formal model. There are the simulations for classical first-order logic that I have focused on here, and there are the Kripke simulations for intuitionist first-order logic, and the simulations for quantified LP and K3, and so on. $\phi \lor \neg \phi$ holds over some classes of models and not others. So it would make more sense to subscript it:

$$\vDash_C \phi \lor \neg \phi$$
$$\nvDash_{Int} \phi \lor \neg \phi$$
$$\nvDash_{K3} \phi \lor \neg \phi$$

And of course, there is no serious disagreement over the truth of any of these three claims, any more than there is disagreement over whether the internal angles of a triangle add up to 180° *in Euclidean geometry*. But there is disagreement over the law of excluded middle. If you want to know whether the unrelativised claim is true, that is, if you are interested in which of these logics is *right*, then you need to go beyond the simulations, asking not just what laws hold over various models, but rather which models we ought to quantify over. And to answer that question we need a reference point outside of the models themselves – something they are supposed to be models of.

5. Conclusion

Logic is the science of truth, so it would be misleading to say that truth does not matter in logic. But as in other sciences, we have to balance the value of accuracy (of the truth of the laws of truth) with other values: simplicity, manageable cognitive or computational load, unification, or elegance. This is both sensible and consistent with a respect for and commitment to truth. We can admit that a simple, strong logic fails to be perfectly accurate, without thereby saying that it is worthless or should be discarded.

If we forget that this is an option, we might suggest that classical logic should be abandoned because we require incomplete predicates to deal with an especially knarly paradox. Or – moving in the other direction – we might reject a complicated subclassical solution to the paradoxes because we are loathe to lose the simplicity and strength of classical logic. Understanding the role of idealisation in logic is helpful because it suggests that neither exclusionary stance is needed: sometimes a false logic – a somewhat imperfect map – is exactly what is required.[25,26]

References

Beall J. (2017). 'There is No Logical Negation: True, False, Both, and Neither.' *The Australasian Journal of Logic* **14**, 1–29.

Borges J.L. (1999). *Collected Fictions*. London: Penguin Classics. translated by Andrew Hurley.

Carroll L. (1894). *Sylvie and Bruno Concluded*. New York: MacMillan and Co.

Cartwright N. (1983). *How the Laws of Physics Lie*. New York: Oxford University Press.

Cook R.T. (2000). *Logic-as-Modelling: A New Perspective on Formalization*. PhD thesis, The Ohio State University.

Etchemendy J. (1999/1990). *On the Concept of Logical Consequence*. Stanford: CSLI.

Frege G. (1956/1918). 'The Thought.' *Mind* **65**(259), 289–311.

Grice P. (1967/1991). *Studies in the Ways of Words*. Cambridge, MA: Harvard University Press.

Hjortland O.T. (2017). 'Anti-Exceptionalism About Logic.' *Philosophical Studies* **174**, 631–58.

Kennedy C. and McNally L. (2005). 'Scale Structure, Degree Modification, and the Semantics of Gradable Predicates.' *Language* **81**(2), 345–81.

Kripke S.A. (1984). 'Outline of a Theory of Truth.' In R.L. Martin (ed.), *Recent Essays on Truth and The Liar Paradox*, pp. 53–81. New York: Oxford University Press.

MacFarlane J. (2009). 'Logical Constants.' In E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*. Palo Alto: Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/entries/logical-constants/

Martin B. and Hjortland O. (2020). 'Logical Predictivism.' *Journal of Philosophical Logic* **50**, 285–318.

Potochnik A. (2017). *Idealization and the Aims of Science*. Chicago: University of Chicago Press.

Russell G.K. (2018). 'Logical Nihilism: Could There Be No Logic?' *Philosophical Issues* **28**, 308–24.

Russell G.K. (2019). 'Deviance and Vice: Strength as a Theoretical Virtue in the Epistemology of Logic.' *Philosophy and Phenomenological Research* **99**(3), 548–63.

Russell G.K. (2022). 'Fancy Loose Talk About Knowledge.' *Inquiry*, **65**(7), 789–820.

Russell G.K. (2023). *Barriers to Entailment: Hume's Law and Other Limits on Logical Consequence*. Oxford: Oxford University Press.

Shaffer M.J. (2012). *Counterfactuals and Scientific Realism*. Basingstoke: Palgrave MacMillan.

Shapiro S. (2001). 'Modelling and Normativity: How Much Revisionism Can We Tolerate?' *Agora: Papeles de Filosofia* **20**(1), 159–73.

Shapiro S. (2009). 'Logical Consequence, Proof, and Model Theories.' In S. Shapiro (ed.), *The Oxford Handbook of Philosophy of Mathematics and Logic*, pp. 651–70. Oxford: Oxford University Press.

Sher G. (1996). 'Did Tarksi Commit "Tarski's Fallacy"?' *Journal of Symbolic Logic* **61**, 182–209.

Soames S. (1999). *Understanding Truth*. New York: Oxford University Press.

Soames S. (2003). *Philosophical Analysis in the Twentieth Century: The Age of Meaning*, Vol. **2**. Princeton, NJ: Princeton University Press.

Tarski A. (1944). 'The Semantic Conception of Truth: And the Foundations of Semantics.' *Philosophy and Phenomenological Research* **4**(3), 341–76.

Tarski A. (1983/1936). 'On the Concept of Logical Consequence.' In J. Corcoran (ed.), *Logic, Semantics and Metamathematics*, 2nd edition, pp. 409–20. Indianapolis: Hackett.

Tarski A. (1986). 'What are Logical Notions?' *History and Philosophy of Logic*, **7**, 143–54. (Transcript of a 1966 talk, ed. J. Corcoran).

Unger P. (1971). 'A Defense of Skepticism.' In F. Dretske and S. Bernecker(eds),, *Knowledge: Readings in Contemporary Epistemology*, pp. 324–38. New York: Oxford University Press.

Varzi A.C. (2002). 'On Logical Relativity.' *Philosophical Issues* **12**, 197–219.